

**UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE QUITO**

**CARRERA:
INGENIERÍA ELECTRÓNICA**

**Trabajo de titulación previo a la obtención del título de:
INGENIERO ELECTRÓNICO**

**TEMA:
CONVERSIÓN DE IMAGEN DEL ESTÁNDAR DE TELEVISIÓN DIGITAL
ISDB-T A ESTEREOSCÓPICA 3D UTILIZANDO SFM**

**AUTOR:
LUIS OMAR SANGOQUIZA CHICAIZA**

**TUTOR:
LUIS GERMÁN OÑATE CADENA**

Quito, julio del 2019

CESIÓN DE DERECHOS DE AUTOR

Yo, Luis Omar Sangoquiza Chicaiza, con documento de identificación N° 1718657917, manifiesto mi voluntad y cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del trabajo de titulación intitulado: “CONVERSIÓN DE IMAGEN DEL ESTÁNDAR DE TELEVISIÓN DIGITAL ISDB-T A ESTEREOSCÓPICA 3D UTILIZANDO SFM”, mismo que ha sido desarrollado para optar por el título de: Ingeniero Electrónico, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En aplicación a lo determinado en la Ley de Propiedad Intelectual, en mi condición de autor me reservo los derechos morales de la obra antes citada. En concordancia, suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Politécnica Salesiana.



.....
Luis Omar Sangoquiza Chicaiza
C.I.: 1718657917

Quito, julio del 2019

DECLARATORIA DE COAUTORIA DEL DOCENTE TUTOR

Yo, declaro que bajo mi dirección y asesoría fue desarrollado el Artículo Académico, “CONVERSIÓN DE IMAGEN DEL ESTÁNDAR DE TELEVISIÓN DIGITAL ISDB-T A ESTEREOSCÓPICA 3D UTILIZANDO SFM” realizado por: Luis Omar Sangoquiza Chicaiza, obteniendo un producto que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana, para ser considerados como trabajo final de titulación.

Quito, julio del 2019



.....

Luis Germán Oñate Cadena

C.I.: 1712157401

Conversión de imagen del estándar de televisión digital ISDB-T a estereoscópica 3D utilizando SFM

Ingeniería Electrónica
Universidad Politécnica Salesiana
Quito, Ecuador
Omar Sangoquiza; Luis Oñate¹

Resumen—La televisión digital busca en los últimos años ofrecer imágenes estereoscópicas en tres dimensiones (3D) sin embargo el problema principal radica en que se necesita un gran número de cámaras para realizar tomas de las imágenes desde distintos ángulos y posiciones añadiendo que esta información sea procesada por un estándar de televisión Digital. Una posible solución es el uso del algoritmo de estructura a partir de movimiento (SFM), este algoritmo permite generar una imagen en 3D a partir de una imagen de televisión digital en dos dimensiones. En este artículo se toma la imagen recibida en un receptor ISDB-T y se construye una imagen en 3D, se analiza mediante la simulación el grado de calidad en la imagen generada en 3D con respecto a la imagen transmitida añadiendo a esto que la imagen depende de los puntos de coincidencia para obtener una imagen en 3D de calidad aceptable, hay que tomar en cuenta el ruido introducido por el canal.

Abstract—Digital TV searches recent years provide stereoscopic images in three dimensions (3D) but the main problem in that a large number of cameras needed for shooting images from different angles and positions adding that this information is processed by a standard of Digital television. A possible solution is the use of the structure algorithm from motion (SFM), this algorithm allows to generate a 3D image from a digital television image in two dimensions. In this article, the image received on an ISDB-T receiver is taken and a 3D image is constructed, the degree of quality in the image generated in 3D with respect to the transmitted image is analyzed by simulation adding to this that the image depends of the points of coincidence to obtain a 3D image of acceptable quality, the noise introduced by the channel must be taken into account.

Palabras Clave: televisión digital, imágenes 3D, SFM, ruido

I. INTRODUCCIÓN

El avance de la tecnología a través de los últimos años ha desarrollado varias técnicas para resolver el problema de obtener el modelo tridimensional de una imagen. Aplicaciones considerablemente usadas tales como: recuperación del patrimonio [1], topografía [2], medicina [3], emplean la técnica SFM para el procesamiento de imágenes.

La técnica SFM es un conjunto de algoritmos que intenta proporcionar información a partir de la identificación de puntos

entre dos o más imágenes estimando la posición de las cámaras que han tomado las imágenes, este algoritmo pretende obtener dicha técnica a partir de una sola cámara, así como la situación tridimensional de los puntos tomados como referencia. [4].

El estudio de la técnica SFM como se detalla en el párrafo anterior, se trata de la más objetiva ya que con el uso de una sola cámara se pretende obtener toda la información sin necesidad de algún elemento extra para poder realizar la conversión [1], al ser un método en el cual se conoce todas las imágenes surge una variante que es la SFM secuencial [5], la cual es posible emplear gracias a que se enfatiza más en la adición de imágenes en el orden que se tomaron, que a su vez han sido procesadas a través del estándar de televisión digital ISDB-T. [6]

Respecto a la variante SFM secuencial en [7] se efectúa una estimación de profundidad mediante métodos directos para determinar la situación tridimensional. La ventaja destacada de este método se enfoca en aprovechar más información de la imagen y obtener reconstrucciones eficientes.

Se implementa [8], el método basado en puntos de interés los cuales contribuyen con más información tal como bordes, esquinas y sus correspondencias en la imagen próxima, gracias a esta información extra se trata de minimizar el error en el momento de la reconstrucción 3D.

Otra implementación [9], tiene algunas observaciones a tener en cuenta para que las imágenes devuelvan buenos resultados, para la obtención de una reconstrucción favorable se debe capturar un mínimo de tres fotografías, pero lo más recomendable es tener entre 10 o 20.

En este artículo se propone el desarrollo de la técnica SFM y la obtención de imágenes a procesar con una sola cámara, adecuando la imagen que es procesada a partir del estándar ISDB-T, con el objetivo de convertirla en tres dimensiones.

El presente artículo se constituye de la siguiente forma: en la Sección I corresponde a la introducción, en la Sección II se describe el algoritmo SFM. En la Sección III se trata acerca del desarrollo del algoritmo SFM y finalmente en la Sección V se evalúa el algoritmo SFM incluyendo la imagen procesada con el estándar de televisión digital ISDB-T.

¹ Luis Omar Sangoquiza Chicaiza, estudiante de la carrera Ingeniería Electrónica mención telecomunicaciones, UPS, Quito
lsangoquizac@est.ups.edu.ec.

Luis German Oñate Cadena, docente de la carrera Ingeniería Electrónica, UPS, Quito, lonate@ups.edu.ec

II. ALGORITMO SFM

El algoritmo parte de las coincidencias de puntos obtenidas entre las dos imágenes durante el emparejamiento. A partir de estas coincidencias, se va a obtener la posición y orientación relativa de las cámaras empleando la matriz fundamental (F) y la matriz esencial (E). La geometría epipolar [10], recoge la geometría proyectiva entre dos vistas, y es encapsulada por la matriz fundamental. Se basa en que, dado un punto en el espacio del que se conoce su posición en una vista, su posición en la otra vista está limitada a una línea epipolar, tal y como se refleja en la figura 1.

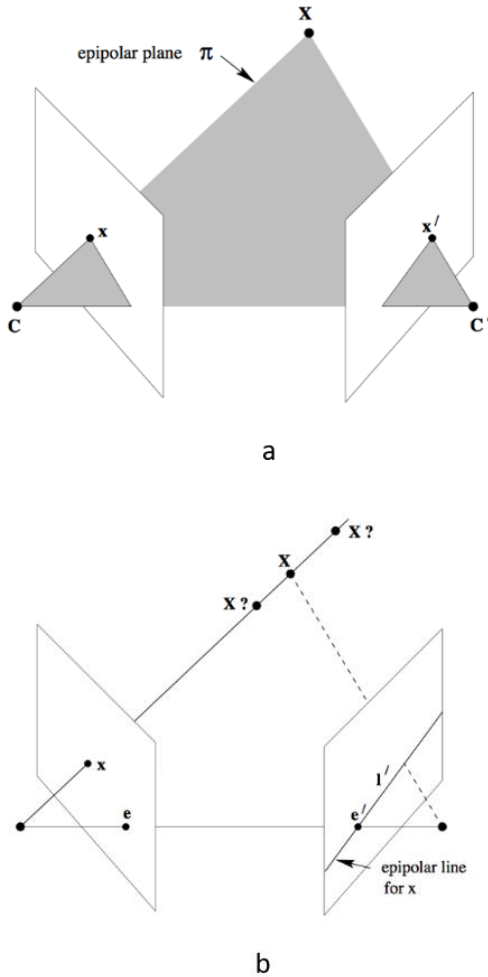


Fig. 1. Geometría epipolar: (a) Plano epipolar (b) Línea epipolar. [10]

Para cada punto que se desea proyectar: el plano epipolar es el que contiene los centros de las cámaras, el punto en el espacio que se desea proyectar y las proyecciones de dicho punto en ambas imágenes. Los puntos epipolares (e y e') son los puntos de las imágenes en los que son cortadas por la línea que une los centros de las cámaras. La línea epipolar (l') es la proyección en la segunda imagen de la línea que proyecta el punto tridimensional en la primera imagen. Por lo tanto, el punto del espacio que se está proyectando, se proyectará en la segunda

imagen en algún punto de la línea epipolar. Se puede ver esta restricción como un mapeo, de forma que para cada punto x en una imagen, existe una línea epipolar l' en la otra vista en la que se encontrará x' . Este mapeo es el que realiza la matriz fundamental, una matriz 3×3 de rango 2 que mapea un punto de un espacio bidimensional (a) a uno unidimensional (la línea epipolar de b). sí x y x' son la proyección del punto X en ambas imágenes en coordenadas homogéneas, al estar x' en l' se cumple: [11]

$$l' = F \cdot x \rightarrow \boxed{x'^T \cdot F \cdot x = 0} = x'^T \cdot l' \quad (1)$$

La relación en el recuadro es fundamental, ya que es la que nos permite computar la matriz fundamental a partir de coincidencias entre puntos de las imágenes. Además, hay que destacar que, en dicha relación las coordenadas de los puntos de la imagen están expresados en píxeles

Para iniciar con el análisis del algoritmo SFM es necesario la calibración de cámara y la selección del tablero. El tablero seleccionado para la calibración debe tener un numero par de cuadros en un eje y un número impar de cuadros en el otro a fin de poder determinar en qué eje se realiza la medida. En la figura 2 se muestra los tipos de tableros los cuales son los idóneos para la calibración (a, b) y el que no puede ser usado para ser calibrado (c). [12]

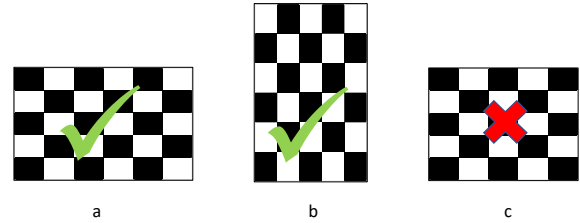


Figura. 2. a) Tablero correcto b) Tablero correcto c) Tablero incorrecto.

Como dato adicional, se debe medir la longitud de la arista de uno de los cuadrados dentro del tablero como se observa en la figura 3 la cual enfoca la dimensión del tamaño entre un cuadro y otro.

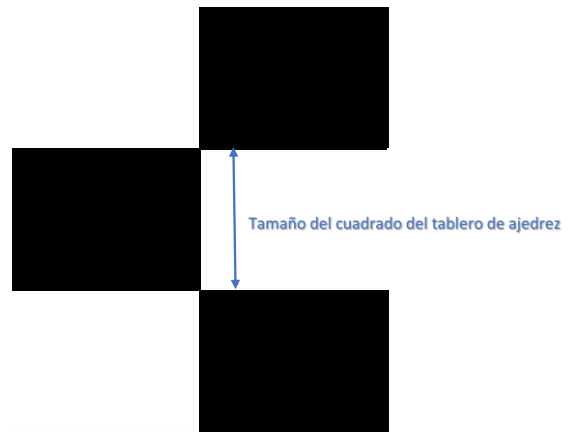


Fig. 3. Dimensión de un cuadro del tablero.

Esta medida puede encontrarse en milímetros, centímetros o pulgadas. Al momento de capturar las imágenes se debe tener en cuenta ciertas consideraciones como que el tablero debe mantenerse enfocado, pero sin usar el auto enfoque incluido a esto el zoom de la cámara tiene que estar estable durante todo el proceso de calibración. Además, si las imágenes se comprimen estas no deben ser usadas. El tablero debe estar a una distancia similar a la distancia que se desea medir en el futuro, no debe estar a más de 45 grados respecto del plano de la cámara, debe ocupar al menos un 20% de la imagen. Finalmente, el algoritmo de calibración requiere por lo menos tres fotografías, pero lo más recomendable es tener un rango de 10 o 20. [12]

El programa analizará cada imagen en búsqueda del tablero en cada una de ellas, el calibrador salta imágenes duplicadas, las posibles razones para no detectarse son una imagen borrosa o un ángulo extremo del patrón. La detección lleva más tiempo con imágenes más grandes y con patrones que contienen una gran cantidad de cuadrados. Finalmente, si el proceso fue exitoso, se muestra una ventana como se observa en la figura 4 la cual muestra la imagen del tablero de ajedrez seleccionada con círculos verdes para indicar los puntos detectados y el cuadrado amarillo indica el origen (0,0). Las flechas X e Y indican la orientación de los ejes del tablero de ajedrez.



Fig. 4. Imágenes sin tablero.

Para la calibración, dependerá del tipo de cámara que se utilice y el uso predeterminado del modelo estándar, que es la opción más adecuada. En la figura 5 se observa la distribución del siguiente contenido: los errores de reproyección y la visualización de parámetros extrínsecos, todo esto calibrado con los valores predeterminados, además se observa un error de reproyección de 1.73 píxeles el cual deberá ser removido para mejorar la estimación.

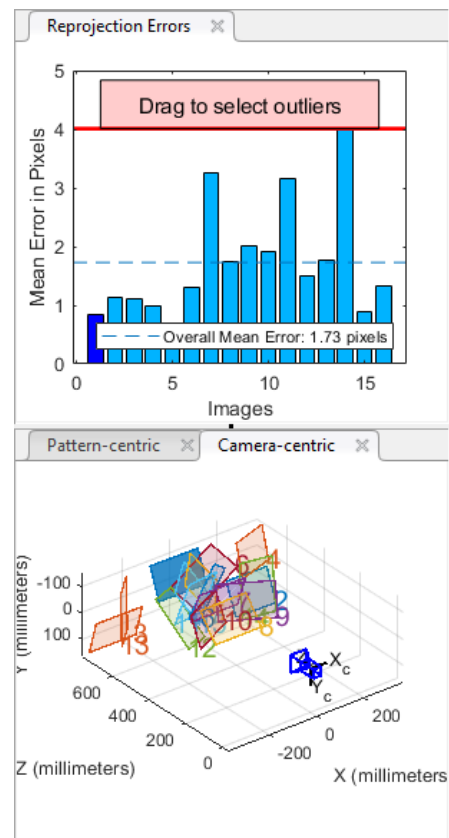


Fig. 5. Calibración con parámetros por defecto.

Para mejorar la estimación en la figura 6 se observa la barra de errores de reproyección con 16 imágenes con su respectivo error en píxeles, a esto se añade la acción de remover la imagen número 14 que es la que contiene el mayor error, por ende, el error baja a 1.63 píxeles.

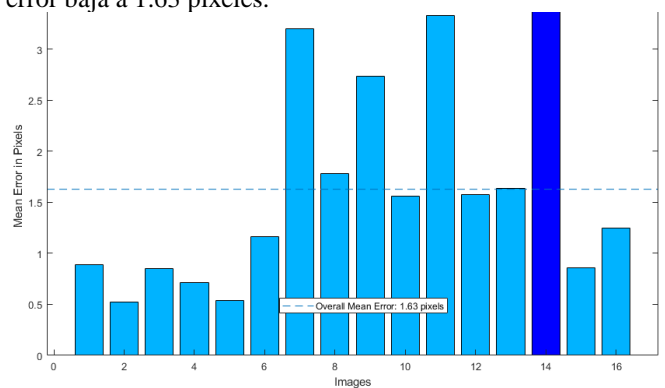


Fig. 6. Error en cada imagen.

En la figura 7 se observa que al remover las imágenes cuyo error fue más alto, se obtiene un error de reproyección de 0.88 píxeles, esto da como resultado un porcentaje de error mejorado de 51%. Además, el gráfico de parámetros extrínsecos en 3D proporciona una vista centrada en la cámara de los patrones y una vista centrada en el patrón de la cámara.

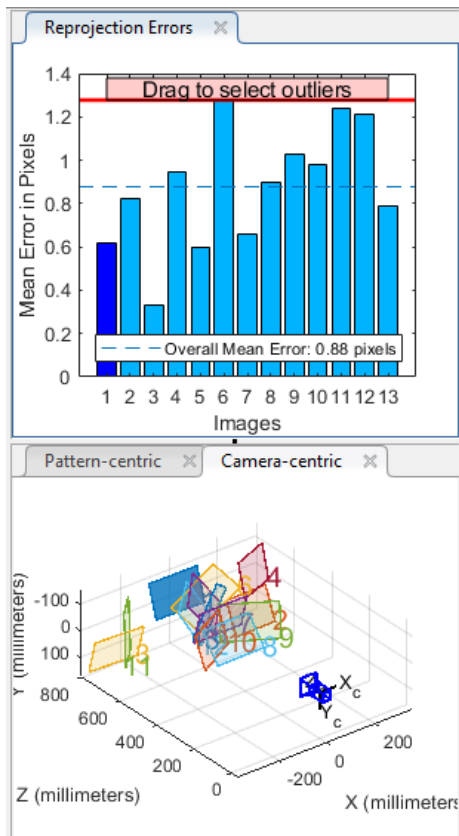


Fig. 7. Resultado final de estimación.

III. DESARROLLO DEL ALGORITMO SFM

La metodología que se sigue para la inicialización del algoritmo SFM es la siguiente [13]:

En la figura 8 se observa el diagrama de flujo resumen del proceso completo que realiza el algoritmo con sus diferentes bloques.

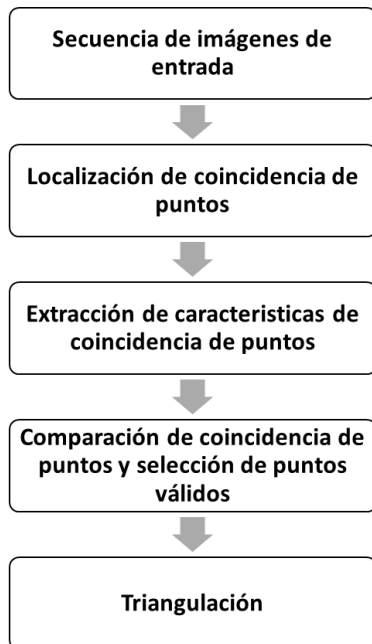


Fig. 8. Diagrama de flujo resumen del algoritmo SFM

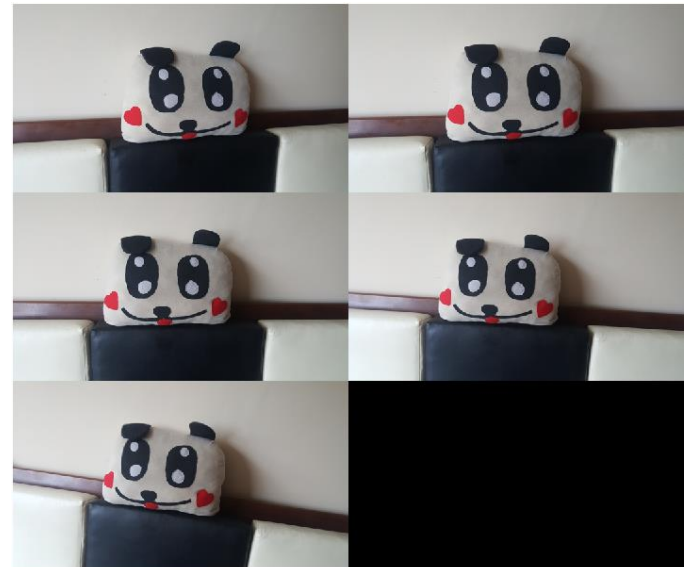


Fig. 9. Imagen de entrada

En la figura 10 se observa los puntos refinados de la cámara además de mostrar las posturas y excluir puntos 3-D ruidosos, especificando el volumen de visualización que no es más que un conjunto de técnicas usadas para mostrar una proyección 2D de un conjunto de datos discretamente muestreados en 3D.

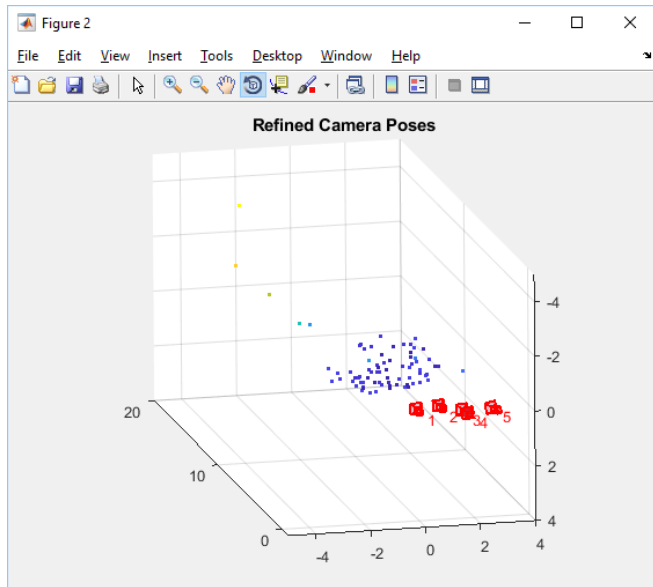


Fig. 10. Posición refinada de cámara

Para finalizar en la figura 11 se observa el resultado final de la reconstrucción de la imagen, además de mostrar las posturas de la cámara y los puntos densos del mundo tridimensional especificando el volumen de visualización.

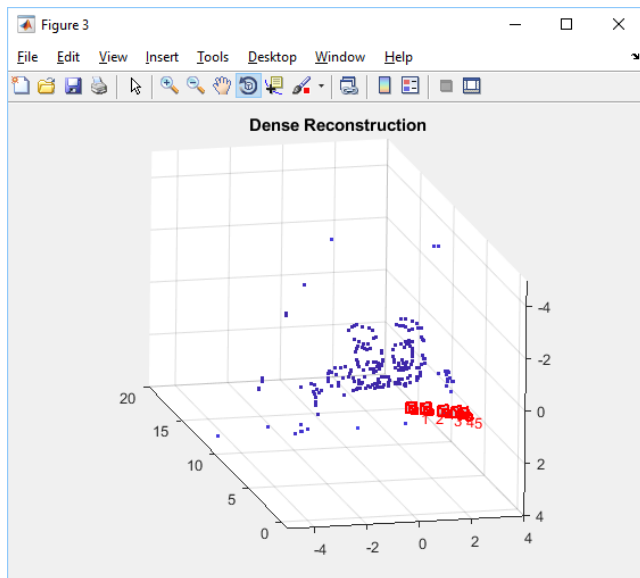


Fig. 11. Resultado final de la reconstrucción

IV. RESULTADOS DE LA EVALUACIÓN DEL ALGORITMO

A continuación, se presentan los resultados de las pruebas ejecutadas al algoritmo SFM. Para evaluarlos se hizo uso de dos simulaciones, en las cuales se va variando el nivel de filtrado para poder diferenciar la intensidad de ruido y la eficiencia en cada imagen convertida.

Las acciones que se llevan a cabo inician con la ejecución del bloque de base de datos en el cual se especificará las imágenes que van a ser procesadas y que fueron previamente

tomadas con la cámara, luego de esto pasan por el diagrama de bloques para el estándar ISDB-T como se observa en la figura 12, este diagrama muestra los bloques de transmisión y recepción para el estándar ISDB-T con modulación OFDM. Los bloques más importantes empleados en la simulación se detallan a continuación: el generador de imagen emplea el bloque Signal From Workspace que es el encargado de importar los datos de la imagen desde el Workspace de Matlab. El Codificador Reed Solomon realiza un proceso de corrección de errores basado en bloques, permitiendo procesar por bloques una cierta cantidad de símbolos de datos. El convertidor de enteros a bits establece un valor de 8, para obtener un total 1632 bits. La modulación de portadora emplea un bloque llamado Matriz interleaver que hará el proceso de entrelazado de bits. Para posteriormente realizar el proceso de Mapping que consta de dos bloques, un Bit to Integer Converter y un Rectangular QAM Modulator. [14]

Finalmente, en el procesamiento de imágenes, se aplica un filtro de mediana cuya ventana es de 3x3 como se observa en la figura 13, además de la secuencia de imágenes del receptor con el estándar ISDB-T ya procesada.

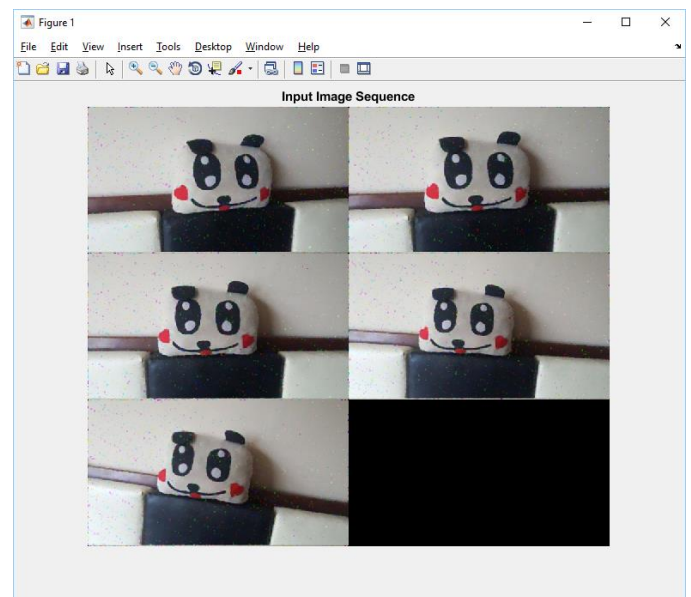


Fig. 13. Secuencia de la imagen de entrada

Para observar los resultados de la reconstrucción 3D, como se observa en la figura 14 la cual contiene la posición refinada de cámaras y en la figura 15 el resultado final de la reconstrucción. Además, se puede apreciar que la reconstrucción posee ciertas falencias, esto es debido a que la imagen al ser procesada bajo el estándar ISDB-T, que para simular la transmisión por un canal multicamino se emplea un subsistema que posee modelos de canales inalámbricos a una línea de 3 caminos (Taps) que altera, distorsiona y retrasa la señal, así como también un bloque "AWGN Channel" que ingresa Ruido Blanco Gaussiano Aditivo. Esta simulación está basada en los tipos de canales SUI (Stanford University Interim), dicha universidad aplicó estos modelos de canales los cuales se basan en 3 tipos de terreno que son los siguientes:

DIAGRAMA DE BLOQUES DE Tx Y Rx DEL ESTÁNDAR ISDB-T CON MODULACIÓN OFDM

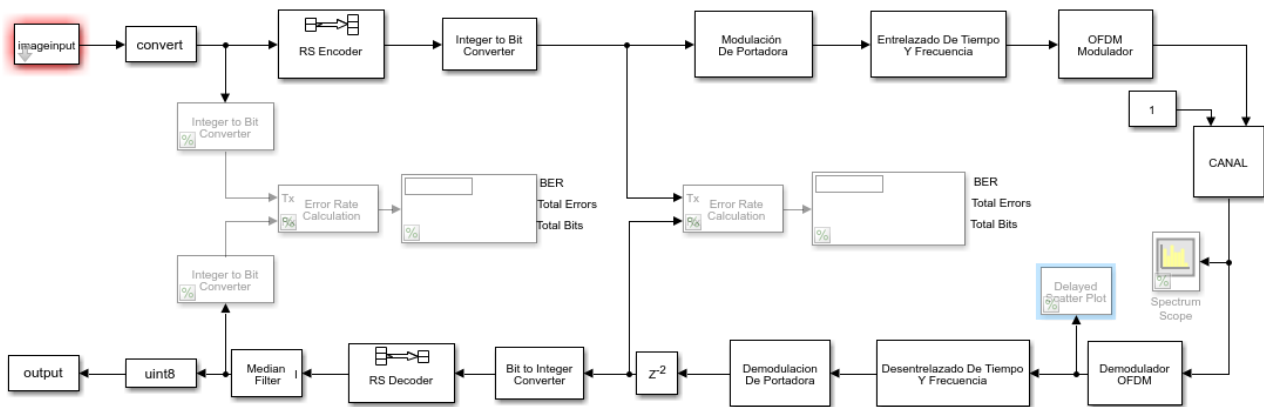


Fig. 12. Diagrama de bloques para el estándar ISDB-T con modulación OFDM elaborado en Simulink de Matlab [14]

Terreno A: medio rural, es el medio más agresivo y corresponde a un entorno montañoso con una vegetación moderada. Terreno B: medio suburbano, es un medio intermedio entre el terreno A y C. Terreno C: medio urbano, es el medio menos agresivo y corresponde a un entorno prácticamente plano con poca densidad de vegetación. [15]

Debido a lo mencionado no se puede encontrar mayores coincidencias de puntos entre imágenes, por tal motivo la reconstrucción no tiene una calidad visibilidad buena.

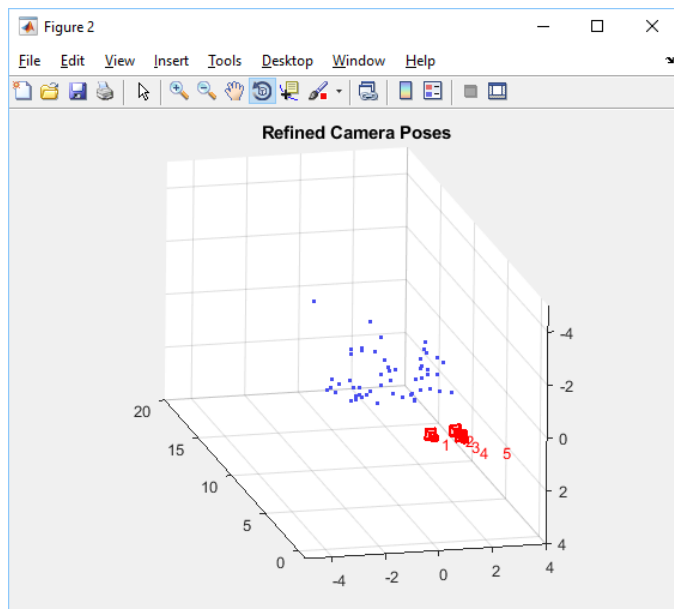


Fig. 14. Posición refinada de cámaras

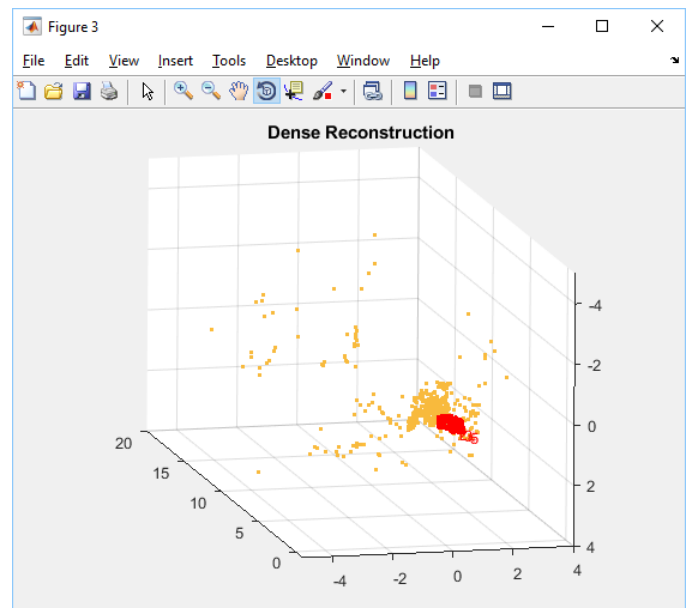


Fig. 15. Resultado final de la reconstrucción

Como se puede observar en la figura 16 la calidad de las imágenes tomadas contiene menos puntos de distorsión, esto debido a la variación del canal hace que desaparezca el ruido y haya más coincidencia de puntos.

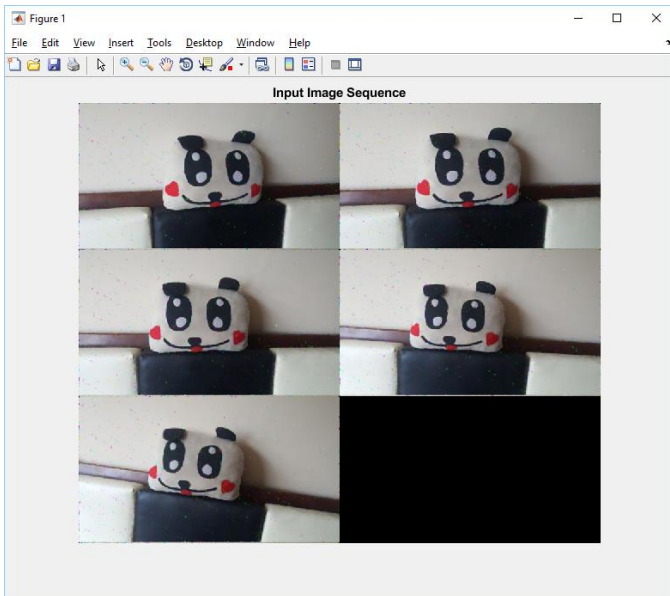


Fig. 16. Secuencia de la imagen de entrada

En la figura 17 y 18 se puede observar la posición refinada de cámaras y el resultado final de la reconstrucción respectivamente, con un nivel de ruido menor haciendo que la coincidencia de más puntos de anotar un leve cambio en la eficiencia al momento de la reconstrucción. Al comparar el algoritmo SFM con un cambio de canal de comunicación se observa que su conversión mejoró apreciándose casi en su totalidad una reconstrucción 3D.

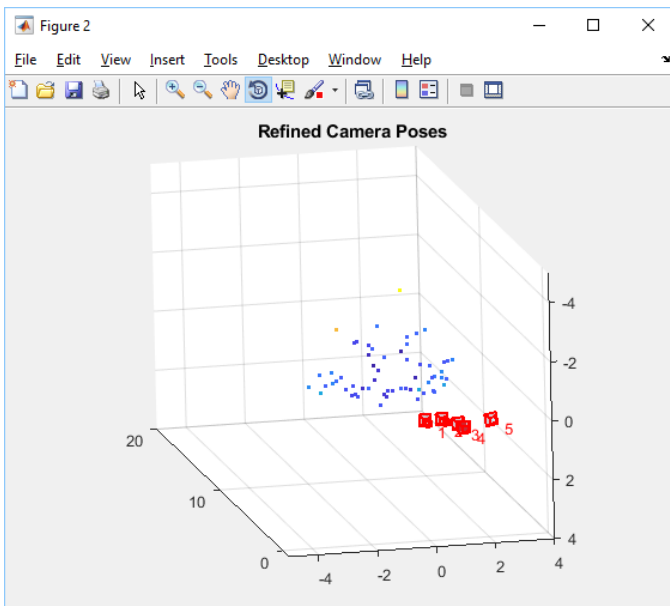


Fig. 17. Posición refinada de cámaras

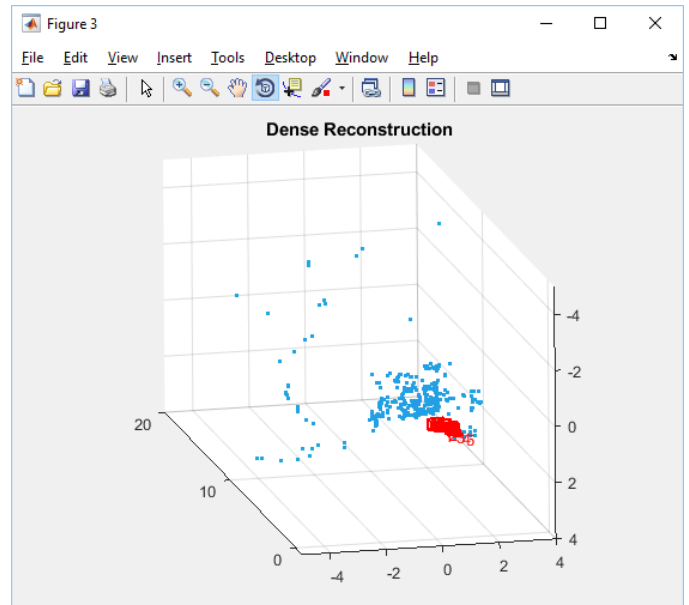


Fig. 18. Resultado final de la reconstrucción

Para finalizar se realizó una comparación entre las figuras 11 y 18 para notar el error medio cuadrado que es una métrica de calidad de imagen aplicando algoritmos de referencia los cuales comparan la imagen de entrada con una imagen de referencia sin distorsión, así se pudo obtener un error de 3,88% al convertirla en 3D, haciendo referencia que la imagen sin el estándar ISDB-T arrojó un valor de 33,87 píxeles y la imagen con el estándar ISDB-T arrojó un valor de 32,58 píxeles. Estos valores están basados en el número de píxeles que posee cada imagen.

V. CONCLUSIONES

De las pruebas realizadas el ruido del canal en el sistema de transmisión ISDB-T que en el Ecuador este proyecto está previsto que trabaje solamente con un canal de 6 MHz influye en que la imagen no pueda ser reconstruida fielmente, la afectación que produce el ruido deteriora la función de coincidencias de puntos con la cual se va obtener la reconstrucción, como solución se puede variar el ruido del canal y el error disminuirá en el cual la imagen reconstruida es más visible.

Al evaluar la eficiencia de conversión de imágenes haciendo una comparación entre estas, se obtuvo un porcentaje de diferencia de 3,88% haciendo menos eficiente la conversión en la cual se utiliza el estándar ISDB-T y el algoritmo SFM.

Para mejorar la reconstrucción 3D de una imagen es necesario utilizar e implementar el uso de cámaras de alta calidad de esta manera las imágenes proporcionadas para procesarse a través del estándar ISDB-T tendrá una mejoría notable encontrando mayores puntos de coincidencia y a su vez mejorando la conversión.

VI. REFERENCIAS

- [1] N. S. R. Snavely, Seitz, Szeliski, «Exploring Photo Collections in 3D,» [En línea]. Available: http://phototour.cs.washington.edu/Photo_Tourism.pdf.
- [2] M. Leotta, «GitHub,» [En línea]. Available: <https://github.com/Kitware/maptk>.
- [3] V. Rizzo, «AUTODESK REMAKE BLOG,» [En línea]. Available: <https://autodeskremake.squarespace.com/blog/2017/1/3/reality-capture-in-biomedicine>.
- [4] C. Abad García, «Repositorio abierto de la Universidad de Cantabria,» Cantabria, 2017.
- [5] D. Bruce, «Stereo and Projective Structure from Motion,» Illinois.
- [6] N. O. PISCIOTTA, «Remultiplexor ISDB-T,» Universidad Blas Pascal, Cordoba, 2014.
- [7] E. Jakob, S. Thomas y C. Danie, «LSD-SLAM: Large-Scale,» Munich, 2014.
- [8] T. Juan, M. Raúl y M. José, «ORB-SLAM: a Real-Time Accurate,» Zaragoza, 2015.
- [9] M. Yi, K. Jana, S. Stefano y S. Shankar. [En línea]. Available: https://www.eecis.udel.edu/~cer/arv/readings/old_mkss.pdf.
- [10] C. D. Singh, «CMSC426 Computer Vision,» [En línea]. Available: <https://cmssc426.github.io/sfm/#featmatch>.
- [11] D. Calderon, «Repositorio Académico de la Universidad de Chile,» 2012. [En línea]. Available: <http://repositorio.uchile.cl/handle/2250/111904>.
- [12] MathWorks. [En línea]. Available: <https://la.mathworks.com/help/vision/ug/single-camera-calibrator-app.html>.
- [13] MathWorks. [En línea]. Available: <https://la.mathworks.com/help/vision/examples/structure-from-motion-from-multiple-views.html>.
- [14] J. Gómez y H. Lapo, Análisis comparativo entre DVB-T e ISDB-T utlizando modulación CC-OFDM, Quito, 2018, p. 24.
- [15] Agarwal; Sameer; Yasutaka, Furukawa; Noah, Snavely, «Building Rome in a Day».